Trees with given degree sequences that have minimal subtrees *

Xiu-Mei Zhang[1,2], Xiao-Dong Zhang[1,†]

[1]Department of Mathematics, Shanghai Jiao Tong University

800 Dongchuan road, Shanghai, 200240, P. R. China

[2]Department of Mathematics,Shanghai sandau University

2727 jinhai road, Shanghai, 201209, P. R. China

Abstract

In this paper, we investigate the structures of an extremal tree which has the minimal number of subtrees in the set of all trees with the given degree sequence of a tree. In particular, the extremal trees must be caterpillar and but in general not unique. Moreover, all extremal trees with a given degree sequence $\pi = (d_1, \cdots, d_5, 1, \cdots, 1)$ have been characterized.

Key words: Tree; subtree; degree sequence; caterpillar;

AMS Classifications: 05C05, 05C30

1 Introduction

Let $T = (V, E)$ be a tree with vertex set $V(T)$ and edge set $E(T)$. vertices of degree 1 of T are called *leaves*. For any two vertices $u, v \in V(T)$, the distance between two vertices u and v, denoted by $d_T(u, v)$ (or $d(u, v)$ for short), is length of the unique path $P_T(u, v)$ joining u and v in T. Then $D(T) = max\{d(u, v)|u, v \in V(T)\}$ is the diameter of tree T. Moreover, we use $N_T(v)$ to indicate the neighbors of vertex v and $d(v) = |N_T(v)|$ is the degree of v. A

*This work is supported by the National Natural Science Foundation of China (No:10971137)

†Corresponding author: Xiao-Dong Zhang (Email: xiaodong@sjtu.edu.cn)

caterpillar is a tree, which has a path, such that every vertex not on the path is adjacent to some vertex on the path.

For a tree $T = (V(T), E(T))$ and $v_1, v_2, \ldots, v_{m-1}, v_m \in V(T)$, let $f_T(v_1, v_2, \ldots, v_{m-1}, v_m)$ denote the number of subtrees of T that contain the vertices $v_1, v_2, \ldots, v_{m-1}, v_m$. In particular, $f_T(v)$ denotes the number of subtrees of T that contain v. Let $\varphi(T)$ denote the number of non-empty subtrees of T. For other terminology and notions, we can follow from [2].

The number of subtrees of a tree has received much attention, since it can reveal some different structures and characterization of a tree. It is well known that the path and the star $K_{1,n-1}$ have the smallest and largest numbers of subtrees among all trees of order n, respectively. Roughly speaking, the less branched the tree is, the smaller the number of subtrees. A observation is that trees with the same maximum degree appear to be clustered together in this order. One may wonder which trees have the largest or smallest Wiener index under the restriction of maximum degree. Kirk and Wang [5] characterized the extremal tree with given a order and maximum vertex degree which have the largest number of subtrees. Recently, Zhang et.al. determined the extremal trees with give a degree sequence that have the largest number of subtrees. For other related results, the authors can be referred to [10, 12]. On the other hand, Heuberger and Prodinger [4] presented formulas to calculate the number of subtrees of extremal trees among binary trees. Yen and Yeh [18] gave a linear-time algorithm to count the subtrees of a tree. Eistenstat and Gordon [3] constructed two non-isomorphic trees that have the same the number of subtrees, which are related to the greedoid Tutte polynomial of a tree. Further an interesting fact is that among above every kind of trees, the extremal one that maximizes the number of subtrees is exactly the one that minimizes some chemical indices such as the well known *Wiener index* (see [17] for details) and vice versa.

Although a counter example has showed in [15] that no 'nice' functional relationship exits between these two concepts, the results are extended to some other kind trees. Such as, recently, the extremal one that maximizes the number of subtrees among trees with a given degree sequence are characterized in [21] and the extremal structures once again coincide with the once found for the Wiener index [16] and [19], respectively. When the extremal one that maximizes the Wiener index with a given degree sequence are istudied in [9] and [20]. Then it is natural to consider the following question.

Problem 1.1 *Given the degree sequence and the number of vertices of a tree, find the lower bound for the number of subtrees, and characterize all extremal trees that attain this bound.*

It will not be a surprise to see that such extremal trees coincide with the ones that attain

the maximal Wiener index. The rest of the paper is organized as follows. In Section 2, we prove that a minimum optimal tree must be a caterpillar. In Section 3, we discuss some properties of the extremal tree with minimal (maximal) number of subtrees among caterpillar trees with given order and degree sequence. In Section 4, the extremal trees with minimal subtrees among given degree sequence $\pi = (d_1, d_2, \cdots, d_n)$, where $d_1 \geq \cdots d_k \geq 2 > d_{k+1} = 1$ and $k \leq 5$ are characterized. Moreover, the extremal minimal trees are not unique.

2 Properties of optimal minimal trees with a given degree sequence

For a nonincreasing sequence of positive integers $\pi = (d_1, d_2, \cdots, d_k, \cdots d_n)$, $d_1 \geq d_2 \geq \cdots \geq d_k \geq 2$ and $d_{k+1} = \cdots = d_n = 1$. If π is the degree sequence of a tree, let $\mathcal{T}_\pi$ denote the set of all trees with π as its degree sequence. For convenience, we refer to trees that maximize (minimize) the number of subtrees as maximum (minimum) optimal. The main result of this section can be stated as follow.

Theorem 2.1 *Let $\pi = (d_1, d_2, \cdots, d_k, \cdots d_n)$ be the degree sequence of a tree with $d_1 \geq d_2 \geq \cdots \geq d_k \geq 2$ and $d_{k+1} = \cdots = d_n = 1$. If T^* is a minimum optimal tree in $\mathcal{T}_\pi$, then T^* must be a caterpillar.*

Proof. Let $T^* \in \mathcal{T}_\pi$ be a minimum optimal tree. If the diameter $D(T^*)$ is equal to 2, then T^* is $K_{1,n-1}$, and also is caterpillar. If $D(T^*) = 3$, then the degree sequence of T^* must be $\pi = (d_1, d_2, 1, \cdots, 1)$ and $d_1 \geq d_2 \geq 2$. It is easy to see that T^* is a caterpillar. Hence we only need to prove the assertion for $D(T^*) \geq 4$ and at least three internal vertices. Let $P = v_0 v_1 v_2 \cdots v_r$ be the longest path in T^*, where $r = D(T^*)$. Then $d(v_0) = d(v_r) = 1$ and $d(v_1), \cdots, d(v_{r-1}) \geq 2$, which implies that there are at least $r-1$ vertices with at least degree 2. So $k \geq r - 1$. Now we have the following claim

Claim: $k = r - 1$.

If $k > r - 1$ then T^* is not a caterpillar. Thus there exists a vertex $y \notin V(P)$ and $2 \leq l \leq r - 2$ such that the edge $yv_l \in E(T)$ and $N_T(y) = \{v_l, x_1, x_2, \cdots, x_s\}$, $s \geq 1$. Moreover, $T - \{v_l v_{l+1}, v_l y\}$ has three connected components, W_1, W_2, W_3 which contain v_l, v_{l+1}, y, respectively. Without loss of generality, we assume that $f_{T_1}(v_l) > f_{T_2}(v_{l+1})$. Further let V_i be the connected component of $T - \{v_{i-1}v_i, v_i v_{i+1}\}$ containing vertex v_i and $a_i = f_{V_i}(v_i)$ for

$l \le i \le r-1$ (for convenience, $a_r = 1$). Moreover, denote $b_l = f_{W_1}(v_l)$ and $a = f_{W_3}(y) > 1$. Then

$$b_l > a_{l+1}(1 + a_{l+2} + a_{l+2}a_{l+3} + \cdots + a_{l+2}a_{l+3}\cdots a_r). \tag{1}$$

Let T' be a tree with degree sequence π obtained from T^* by deleting the edges $yx_1, yx_2, \cdots, yx_s$ in T^* and adding the edges $v_rx_1, v_rx_2, \cdots, v_rx_s$. Obviously, $T' \in \mathcal{T}_\pi$. Let W_3' be the connected component of $T' - \{v_{r-1}v_r\}$ containing vertex v_r. Then W_3' is isomorphic to W_3 and $f_{W_3'}(v_r) = f_{W_3}(y) = a$. Clearly, the number of subtrees of T^* with containing v_r, y is equal to the number of subtrees of T' with containing v_r, y. The number of of subtrees of T^* without containing v_r, y is equal to the number of subtrees of T' without containing v_r, y. The number of of subtrees of T^* with containing y and no containing v_r is equal to $a(1+b_l+b_la_{l+1}+\cdots+b_la_{l+1}\cdots a_{r-1})$, while the number of of subtrees of T' with containing y and no containing v_r is equal to $1+b_l+b_la_{l+1}+\cdots+b_la_{l+1}\cdots a_{r-1}$. The number of of subtrees of T^* with containing v_r and no containing y is equal to $1+a_{r-1}+a_{r-1}a_{r-2}+\cdots+a_{r-1}a_{r-2}\cdots a_{l+1}b_l$, while the number of of subtrees of T' with containing v_r and no containing y is equal to $a(1 + a_{r-1} + a_{r-1}a_{r-2} + \cdots + a_{r-1}a_{r-2}\cdots a_{l+1}b_l)$. Hence by equation (1) and $a > 1$,

$$\begin{aligned}\varphi(T') - \varphi(T^*) &= (1-a)[b_l + b_la_{l+1} + b_la_{l+1}a_{l+2} + \cdots + b_la_{l+1}a_{l+2}\cdots a_{r-2} \\ &\quad -a_{r-1} - a_{r-1}a_{r-2} - \cdots - a_{r-1}a_{r-2}\cdots\ a_{l+1}] \\ &< 0.\end{aligned}$$

It contradicts to T^* being an minimum optimal tree. Hence the claim holds and T^* is a caterpillar. ■

3 Properties of optimal trees among caterpillars with a given degree sequence

In this section, we study some properties of optimal minimal (maximal) trees in the set of all caterpillars for a given degree sequence, since an optimal minimal tree in the set of all trees with a given degree sequence must be caterpillar. For graphic sequence $\pi = (d_1, d_2, \cdots, d_k, \cdots d_n)$ of a tree with $d_1 \ge d_2 \ge \cdots \ge d_k \ge 2$ and $d_{k+1} = \cdots = d_n = 1$ with $k \ge 2$, let

$$C_\pi = \{T : \ \ T \text{ is a caterpillar with degree sequence } \pi\}.$$

If $(y_1, y_2, \cdots, y_k)$ is a permutation of $(d_1-2, d_2-2, \cdots, d_k-2)$, then the caterpillar $C(y_1, \cdots, y_k)$ is obtained from a path $v_0v_1v_2\cdots v_kv_{k+1}$ by adding $y_1, \cdots, y_k$ pendent edges at $v_1, \cdots, v_k$,

respectively. Clearly, $C(y_1, \cdots, y_k) \in C_\pi$. Conversely, for any $T \in C_\pi$, T can be obtained in this way. Moreover, let V_j, $V_{\geq j}$ and $V_{\leq j}$ denote the connected component of $C(y_1, y_2, \cdots, y_k)$ containing v_j after deleting the two edges $v_{j-1}v_j$ and v_jv_{j+1}, the edge $v_{j-1}v_j$, and the edge v_jv_{j+1}, respectively, for $j = 1, \cdots, k$. For convenience, let $V_0 = V_{\leq 0} = \{v_0\}$ and $V_{k+1} = V_{\geq k+1} = \{v_{k+1}\}$.

Lemma 3.1 *Let T be a tree $C(y_1, \cdots, y_k)$ in C_π with the spine $v_0v_1 \cdots v_{k+1}$. If there exists a $2 \leq p \leq k-1$ such that $f_{V_{p-i}}(v_{p-i}) \geq f_{V_{p+i}}(v_{p+i})$ for $i = 1, \cdots, q$ and $q \leq min\{k-p, p-1\}$ with at least one strict inequality and $f_{V_{\leq p-q-1}}(v_{p-q-1}) > f_{V_{\geq p+q+1}}(v_{p+q+1})$, then there exists a caterpillar $T_1 \in C_\pi$ such that*

$$\varphi(T_1) < \varphi(T).$$

Proof. Let W be the connected component of T by deleting the two edges $v_{p-q-1}v_{p-q}$ and $v_{p+q}v_{p+q+1}$ and containing vertices v_{p-q} and v_{p+q}. Let X be obtained from the $V_{\leq p-q}$ by adding the edge $v_{p-q-1}v_{p-q}$ and let Y be obtained from the $V_{\geq p+q+1}$ by adding the edge $v_{p+q}v_{p+q+1}$. Then $f_X(v_{p-q}) > f_Y(v_{p+q})$. Further, by Lemma 3.1 in [5], $f_W(v_{p-q}) > f_W(v_{p+q})$. Now let T_1 be the caterpillar from T by deleting two edges $v_{p-q-1}v_{p-q}$ and $v_{p+q}v_{p+q+1}$ and adding two edges $v_{p-q-1}v_{p+q}$ and $v_{p+q}v_{p-q-1}$. Then $T_1 \in C_\pi$. By Lemma 3.2 in [5], $\varphi(T_1) < \varphi(T)$. Hence the assertion holds. ■

Similarly, we can prove the following assertion by the same method and omit the detail.

Lemma 3.2 *Let T be a tree $C(y_1, \cdots, y_k)$ in C_π with the spine $v_0v_1 \cdots v_{k+1}$. If there exists a $1 \leq p \leq k-1$ such that $f_{V_{p-i}}(v_{p-i}) \geq f_{V_{p+i+1}}(v_{p+i+1})$ for $i = 0, 1, \cdots, q$ and $q \leq min\{k-p-1, p-1\}$ with at least one strict inequality and $f_{V_{\leq p-q-1}}(v_{p-q-1}) > f_{V_{\geq p+q+2}}(v_{p+q+2})$, then there exists a caterpillar $T_1 \in C_\pi$ such that*

$$\varphi(T_1) < \varphi(T).$$

It follows from Lemmas 3.1 and 3.2 that we have got a property of an optimally minimal caterpillar tree in C_π.

Corollary 3.3 *Let T be a minimum optional tree $C(z_1, \cdots, z_k)$ with the spine $v_0v_1v_2 \cdots v_kv_{k+1}$ in C_π.*

(i). If there exists a $2 \leq p \leq k-1$ such that $f_{V_{p-i}}(v_{p-i}) \geq$ (or $\leq$) $f_{V_{p+i}}(v_{p+i})$ for $i = 1, \cdots, q$ and $q \leq min\{k-p, p-1\}$ with at least one strict inequality, then

$$f_{V_{\leq p-q-1}}(v_{p-q-1}) \leq \text{ (or } \geq) \ f_{V_{\geq p+q+1}}(v_{p+q+1}). \tag{2}$$

(ii). If there exists a $1 \le p \le k-1$ *such that* $f_{V_{p-i}}(v_{p-i}) \geq$ (or $\leq$) $f_{V_{p+i+1}}(v_{p+i+1})$ $i = 0, \cdots, q$ *and* $q \le min\{k-p-1, p-1\}$ *with at least one strict inequality, then*

$$f_{V_{\le p-q-1}}(v_{p-q-1}) \le \text{ (or } \ge) \ f_{V_{\ge p+q+2}}(v_{p+q+2}). \tag{3}$$

Further, we present another property of a minimal optimal tree in C_π.

Theorem 3.4 *Let* $\pi = (d_1, d_2, \cdots, d_k, \cdots d_n)$ *be the degree sequence of a tree with* $d_1 \ge d_2 \ge \cdots \ge d_k \ge 2$ *and* $d_{k+1} = \cdots = d_n = 1$, *where* $k \ge 3$. *If* $C(z_1, z_2, \cdots, z_k)$ *is a minimal optional caterpillar in* C_π *with* $z_1 \ge z_k$, *then there exists a positive integer number* $1 \le t \le k-1$ *such that*

$$z_1 \ge z_2 \cdots \ge z_{t-1} > z_t = d_k - 2$$

and

$$z_t \le z_{t+1} \cdots \le z_k.$$

Proof.. We consider the following three cases.

Case 1: $d_1 = \cdots = d_k$. Clearly, the assertion holds.

Case 2: $d_1 > d_2 = \cdots = d_k$. Suppose (for contradiction) that there exists $2 \le l \le k-1$ such that $z_l = d_1 - 2$. If l is odd, let $2 \le p = \frac{l+1}{2} \le k-1$ and $q = \frac{l-1}{2}$. Then $f_{V_{p-i}}(v_{p-i}) = f_{V_{p+i}}(v_{p+i})$ for $i = 1, \cdots, q-1$ and $f_{V_{p-q}}(v_{p-q}) = 2^{d_2-2} < 2^{d_1-2} = f_{V_{p+q}}(v_{p+1})$. Hence by (i) of Corollary 3.3, we have $f_{V_{\le p-q-1}}(v_{p-q-1}) \ge f_{V_{\ge p+q+1}}(v_{p+q+1})$, which implies $1 = f_{V_0}(v_0) \ge f_{V_{\ge p+q+1}}(v_{p+q+1}) = f_{V_{\ge l=1}}(v_{l+1})$. It is a contradiction. If l is even, let $p = \frac{l}{2}$ and $q = \frac{l}{2} - 1$. Similarly, by (ii) of Corollary 3.3, we have $1 = f_{V_0}(v_0) \ge f_{V_{\ge l=1}}(v_{l+1})$. It is a contradiction. So the assertion holds.

Case 3: $d_2 > d_k$. We have the following Claim: $z_k > d_k - 2$. In fact, suppose that $z_k = d_k - 2$. Then there exists a $3 \le s \le k$ such that $z_{s-1} > z_s = \cdots = z_k = d_k - 2$. If $k+s$ is odd, let $p = \frac{k+s-1}{2}$ and $q = \frac{k-s+1}{2}$. Then $f_{V_{p-i}}(v_{p-i}) = f_{V_{p+i}}(v_{p+i}$ for $i = 1, \cdots, q-1$ and

$$f_{V_{p-q}}(v_{p-q}) = f_{V_{s-1}}(v_{s-1}) = 2^{z_{s-1}} > 2^{z_k} = f_{V_k}(v_k) = f_{V_{p+q}}(v_{p+q}).$$

Then by (i) of Corollary 3.3, we have

$$f_{V_{\le s-2}}(v_{s-2}) = f_{V_{\le p-q-1}}(v_{p-q-q}) > f_{V_{\ge p+q+1}}(v_{p+q+1}) = f_{V_{k+1}}(v_{k+1}) = 1.$$

It is a contradiction. If $k+s$ is even, by similar method and applying (ii) of Corollary 3.3, we also get the contradiction. Hence the Claim holds.

Further, there exists two integers $2 \leq t \leq l \leq k-1$ such that $z_{t-1} > z_t = d_k - 2$ and $z_t = \cdots = z_l < z_{l+1}$. Then $f_{V_{t-1}}(v_{t-1}) = 2^{z_{t-1}} > 2^{z_t} = f_{V_t}(v_t)$. Hence by (i) of Corollary 3.3, $f_{V_{\leq t-2}}(v_{t-2}) \leq f_{V_{\geq t+1}}(v_{t+1})$. Therefore, for any $1 \leq j \leq t-2$ we have

$$f_{V_{\leq j-1}}(v_{j-1}) < f_{V_{\leq t-2}}(v_{t-2}) \leq f_{V_{\geq t+1}}(v_{t+1}) \leq f_{V_{\leq j+2}}(v_{j+2}).$$

Hence by (i) of Corollary 3.3, we have $f_{V_j}(v_j) \geq f_{V_{j+1}}(v_{j+1})$, i.e., $2^{z_j} \geq 2^{z_{j+1}}$. So $z_1 \geq z_2 \geq \cdots \geq z_{t-1} > z_t$.

Since $z_l < z_{l+1}$, we have $f_{V_l}(v_l) < f_{V_{l+1}}(v_{l+1})$. By (i) of Corollary 3.3, $f_{V_{\leq l-1}}(v_{l-1}) > f_{V_{\geq l+2}}(v_{l+2})$. Then for any $l+1 \leq j \leq k-1$, we have

$$f_{V_{\leq j-1}}(v_{j-1}) \geq f_{V_{\leq l-1}}(v_{l-1}) \geq f_{V_{\geq l+2}}(v_{l+2}) > f_{V_{\leq j+2}}(v_{j+2}).$$

Hence by (i) of Corollary 3.3, we have $f_{V_j}(v_j) \leq f_{V_{j+1}}(v_{j+1})$, i.e., $2^{z_j} \leq 2^{z_{j+1}}$. So $z_l < z_{l+1} \leq \cdots \leq z_k$. We finish our proof. ■

Similarly, there is a property of a maximum optional caterpillar with a given degree sequence.

Theorem 3.5 *Let $\pi = (d_1, d_2, \cdots, d_k, \cdots d_n)$ be the degree sequence of a tree with $d_1 \geq d_2 \geq \cdots \geq d_k \geq 2$ and $d_{k+1} = \cdots = d_n = 1$, where $k \geq 3$. If $C(z_1, z_2, \cdots, z_k)$ is a maximal optional caterpillar in C_π, then there exists an integer $1 \leq t \leq k-1$ such that*

$$z_1 \leq z_2 \cdots \leq z_{t-1} < z_t$$

and

$$z_t \geq z_{t+1} \cdots \geq z_k.$$

Proof. The proof is similar to that of Theorem 3.4 and omitted. ■

4 The optimal minimal trees with many leaves

In this section, for a given degree sequence $\pi = (d_1, d_2, \cdots, d_k, \cdots d_n)$ with at least $n-5$ leaves, we give the minimal optimal trees with the minimum number of subtrees in $\mathcal{T}_\pi$. Moreover, the minimal optimal trees may be not unique.

Theorem 4.1 *Let $\pi = (d_1, d_2, \cdots, d_k, \cdots, d_n)$ be tree degree sequence with $n-k$ leaves for $2 \le k \le 4$. Then the minimal tree in $\mathcal{T}_\pi$ is unique. In other words,*

(i). If $k = 2$, then $\varphi(T) = 2^{n-2} + 2^{d_1-1} + 2^{d_2-1} + n - 2$ for any $T \in \mathcal{T}_\pi$.

(ii). If $k = 3$, then for any $T \in \mathcal{T}_\pi$,

$$\varphi(T) \ge \varphi(C(d_1-2, d_3-2, d_2-2)) = n - 3 + 2^{d_1-1} + 2^{d_2-1} + 2^{d_3-2} + 2^{d_1+d_3-3} + 2^{d_3+d_2-3} + 2^{n-3}.$$

with equality if and if T is the caterpillar $C(d_1-2, d_3-2, d_2-2)$.

(iii). If $k = 4$, then

$$\begin{aligned}\varphi(T) &\ge \varphi(C(d_1-2, d_4-2, d_3-2, d_2-2)) \\ &= n - 4 + 2^{d_1-1} + 2^{d_2-1} + 2^{d_3-2} + 2^{d_4-2} + 2^{d_1+d_4-3} + 2^{d_3+d_4-4} + 2^{d_3+d_2-3} \\ &\quad +2^{d_1+d_4+d_3-5} + 2^{d_2+d_3+d_4-5} + 2^{n-4}\end{aligned}$$

with equality if and only if T is the caterpillar $C(d_1-2, d_4-2, d_3-2, d_2-2)$.

Proof. (i). $k = 2$. Then T must be $C(d_1-2, d_2-2)$ or $C(d_2-1, d_1-2)$, it is to see that

$$\varphi(C(d_1-2, d_2-2)) = \varphi(C(d_2-1, d_1-2)) = 2^{n-2} + 2^{d_1-1} + 2^{d_2-1} + n - 2.$$

(ii). $k = 3$. Then T must be one of $C(d_1-2, d_2-2, d_3-2)$,$C(d_1-2, d_3-2, d_2-2)$ and $C(d_2-2, d_1-2, d_3-2)$. By Theorem **??**,

$$\varphi(T) \ge \varphi(C(d_1-2, d_2-2, d_3-2))$$

with equality if and only if T is $C(d_1-2, d_3-2, d_2-2))$ and

$$\varphi(C(d_1-2, d_3-2, d_2-2)) = n - 3 + 2^{d_1-1} + 2^{d_2-1} + 2^{d_3-2} + 2^{d_1+d_3-3} + 2^{d_3+d_2-3} + 2^{n-3}.$$

(iii). $k = 4$. Let T^* be any minimal optimal tree in the set of all trees with $\pi = (d_1, d_2, d_3, d_4, 1, \cdots, 1)$ and $d_4 \ge 2$. By Theorem 2.1 and 3.4, T^* must be caterpillar and be one of $C(d_1-2, d_2-2, d_4-2, d_3-2)$, $C(d_1-2, d_3-2, d_4-2, d_2-2)$ and $C(d_1-2, d_4-2, d_3-2, d_2-2)$. Further, by lemma 3.1, T^* must be $C(d_1-2, d_4-2, d_3-2, d_2-2)$ and

$$\varphi(T^*) = n - 4 + 2^{d_1-1} + 2^{d_2-1} + 2^{d_3-2} + 2^{d_4-2} + 2^{d_1+d_4-3} + 2^{d_3+d_4-4} + 2^{d_3+d_2-3} + 2^{d_1+d_4+d_3-5} + 2^{d_2+d_3+d_4-5} + 2^{n-4}.$$

This completes the proof. ■

Theorem 4.2 *Let $\pi = (d_1, d_2, \cdots, d_5, 1, \cdots, 1)$ be tree degree sequence with $n-5$ leaves.*

(i). If $2^{d_1} > 2^{d_3-1}(1+2^{d_2-1})$ and $d_4 \neq d_5$, then there is exact one minimal optional tree $C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)$ in $\mathcal{T}_\pi$

(ii). If $2^{d_1} = 2^{d_3-1}(1+2^{d_2-1})$ or $d_4 = d_5$, then there are exact two minimal optional trees $C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)$ and $C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)$ in $\mathcal{T}_\pi$.

(iii). If $2^{d_1} < 2^{d_3-1}(1+2^{d_2-1})$ and $d_4 \neq d_5$, there is exact one minimal optional tree $C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)$ in $\mathcal{T}_\pi$.

Proof. Let T^* be any minimal optimal tree in $\mathcal{T}_\pi$. By Theorem 2.1 and 3.4, T^* must be caterpillar and $C(d_1-2, x_2, x_3, x_4, d_2-2)$, where (x_2, x_3, x_4) is a permutation of (d_3-2, d_4-2, d_5-2).

Further, by Lemma 3.1 , T^* must be $C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)$ or $C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)$. Moreover,

$$\begin{aligned} & \varphi(C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)) - \varphi(C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)) \\ = \; & 2^{d_5-2} + 2^{d_1+d_5-3} + 2^{d_4-2} + 2^{d_3+d_4-4} + 2^{d_2+d_3+d_5-3} \\ & -[2^{d_4-2} + 2^{d_1+d_4-3} + 2^{d_5-2} + 2^{d_3+d_5-4} + 2^{d_2+d_3+d_4-3}] \\ = \; & (2^{d_5-2} - 2^{d_4-2})[2^{d_1} - 2^{d_3-1}(1+2^{d_2-1})]. \end{aligned}$$

If $2^{d_1} > 2^{d_3-1}(1+2^{d_2-1})$ and $d_4 \neq d_5$, then

$$\varphi(C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)) - \varphi(C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)) < 0.$$

Hence (i)holds.

If $2^{d_1} < 2^{d_3-1}(1+2^{d_2-1})$ and $d_4 \neq d_5$, then

$$\varphi(C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)) - \varphi(C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)) > 0.$$

Hence (iii)holds.

If $2^{d_1} = 2^{d_3-1}(1+2^{d_2-1})$ or $d_4 = d_5$, then

$$\varphi(C(d_1-2, d_5-2, d_4-2, d_3-2, d_2-2)) = \varphi(C(d_1-2, d_4-2, d_5-2, d_3-2, d_2-2)).$$

Hence (ii) holds. ■

Remark From Theorem 4.2, we can see that the minimal optimal trees depend on the values of all components of the tree degree sequences and not unique, while the maximal optimal tree is unique for a given tree degree sequence. It illustrates that it is difficult to character the minimal optimal trees for a given degree sequence of a tree.

References

[1] A. Agra, L. Gouveia and C. Requejo, Extended formulations for the cardinality constrained subtree of a tree problem, Operations Research Letters, 37 (2009) 192–196.

[2] J. A. Bondy and U.S.R. Murty, Graph theory with applications, Macmillan Press,New York, 1976.

[3] D. Eisenstat and G. Gordon, Non-isomorphic caterpillars with identical subtree data, Discrete Math. 306 (2006) 827–830.

[4] C. Heuberger and H. Prodinger, On α-greedy expansions of numbers, Adv. in Appl. Math. 38 (2007) 505–525.

[5] R. Kirk and H. Wang, Largest number of subtrees of trees with a given maximum degree, SIAM J. Discrete Math. 22 (2008) 985–995.

[6] B. Knudsen, Optimal multiple parsimony alignment with affine gap cost using a phylogenetic tree, in Lecture Notes in Bioinformatics, Vol.2812, Springer-Verlag, 2003, pp. 433–446.

[7] A. Meir and J. W. Moon, On subtrees of certain families of rooted trees, Ars Combin. 16 (1983) 305-318.

[8] The On-Line Encyclopedia of Integer Sequences, A092781, http://www.research.att.com/ njas/sequences.

[9] R. Shi, The average distance of trees, Systems Science and Mathematical Sciences, 6 (1)(1993) 18–24.

[10] L. A. Székely and H. Wang, On subtrees of trees, Advances in Applied Mathematics, 34 (2005) 138–155.

[11] L. A. Székely and H. Wang, Binary trees with the largest number of subtrees with at least one leaf, Congr. Numer. 177 (2005) 147–169.

[12] L. A. Székely and H. Wang, Binary trees with the largest number of subtrees, Discrete Applied Mathematics, 155 (2007) 374–385.

[13] L. A. Székely, H. Wang, and T. Y. Wu, The sum of the distances between the leaves of a tree and the ”semi-regular” property, Discrete Mathematics, 311 (2011) 1197-1203.

[14] A. Vince and H. Wang, The average order of a subtree of a tree, Journal of Combintorial Theory, Ser. B, 100 (2010) 161-170.

[15] S. G. Wagner, Correlation of graph-theoretical indeces, SIAM J. Discrete Math. 21 (2007) 33–46.

[16] H. Wang, The extremal values of the Wiener index of a tree with given degree sequence, Discrete Applied Mathematics, 156 (2008) 2647–2654.

[17] Wiener, H., Structural determination of paraffin boiling poiling points, J. Amer. Chem. Soc. 69 (1947) 17-20.

[18] W.-G. Yan and Y.-N. Yeh, Enumeration of subtrees of trees, Theoretical Comuter Science, 369 (2006) 256–268.

[19] X.-D. Zhang, Q.-Y. Xiang, L.-Q. Xu, and R.-Y. Pan, The Wiener Index of Trees with Given Degree Sequences, MATCH Commun.Math.Comput.Chem. 60 (2008) 623–644.

[20] X.-D. Zhang, Y. Liu and M.-X. Han, Maximum Wiener Index of Trees with Given Degree Sequence, MATCH Commun.Math.Comput.Chem. 64 (2010) 661–682.

[21] X.-M. Zhang, X.-D. Zhang, D. Gray and H. Wang, The Number of Subtrees with Given Degree Sequence, submitted for publicatin.